DOCUMENT RESUME

ED 310 157 TM 013 773

AUTHOR Higgins, Raymond E.
TITLE Approaches to Evaluation in Art Education at the State Level.
PUB DATE 89
NOTE 12p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Art Education; Curriculum Development; Curriculum Evaluation; *Educational Assessment; Elementary Secondary Education; *Evaluation Methods; Outcomes of Education; *Program Evaluation; Standardized Tests; State Legislation; State Programs; *Statewide Planning
IDENTIFIERS Minnesota State Department of Education

ABSTRACT

This study reviewed the context of and activities involved in state-level evaluation of elementary/secondary art education in Minnesota. Four main areas reviewed include: (1) legislation impacting all state-level assessment decisions; (2) curriculum development initiatives as they relate to assessment; (3) a specific state-level assessment project; and (4) the status of the statewide assessment program. State legislation upon which art education evaluations are based is referred to as the "PER Law", which stands for Planning, Evaluating, and Reporting. The law requires the Minnesota Department of Education to identify essential learner outcomes that must be incorporated into each district's educational plan; the result is a degree of curriculum standardization. Additional statewide standardization is provided by standardized testing, currently implemented via a multiple-choice, paper-and-pencil format. Other possible sources of information about student skill and knowledge, such as interactive video techniques, self-reports of students, observational techniques, and reflective interviews, are under consideration. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED310157

Approaches to Evaluation in Art Education at the State Level

Raymond E. Higgins, Ph.D.
Art Education Specialist
Minnesota Department of Education

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
RAYMOND E. HIGGINS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

There is an increasing call for state departments of education to develop comprehensive testing and evaluation procedures at both program and individual student levels (National Endowment, 1988). With increased interest in accountability in all subject areas, great emphasis must be placed on developing valid, reliable, and meaningful evaluation procedures. Curriculum and assessment specialists at the Minnesota Department of Education are actively engaged in the development of evaluation procedures for all discipline areas (Minnesota Statewide). Several different approaches to evaluation are presently being investigated.

However, before addressing the nature of state level assessment activities in Minnesota it is necessary to provide a comprehensive picture of the relationship of assessment to curriculum development in the state. This will be accomplished by reviewing four main areas: 1) legislation which impacts all state level assessment decisions; 2) curriculum development initiatives as they relate to assessment; 3) a specific state level assessment project; and 4) the status of the statewide asssessment program.

Legislation

In 1976 the Minnesota state legislature enacted a significant piece of legislation which continues to affect all of K-12 education in the state. This legislation is referred to as the PER law, which stands for Planning, Evaluating, and Reporting. Legislators felt that improvement of local educational programs would best be facilitated by well-defined efforts in curriculum planning and evaluation. Originally the PER legislation merely encouraged local school districts to identify and implement a process to plan and evaluate their educational programs. In 1978-79 the process of planning and evaluating became mandatory, hence the significance of the 'R' in the PER acronym. Now it is required that all districts in the state report a written policy which describes the process they have adopted for curriculum planning and evaluation at the local level (Randall, 1989).

Over the past decade the PER legislation has been amended many times, in each case adding further stipulations and requirements to the original law. Many of these changes directed

TM013773

BEST COPY AVAILABLE

school districts to show further evidence of efforts to develop more comprehensive and effective programs which also meet State Board of Education requirements. Specific wording was added requiring an evaluation of the testing program in each district. To help oversee and coordinate the PER process facilitators were identified and funded by the state at regional locations, the Edcuational Cooperative Service Units. As the law became increasingly prescriptive allowances were made for state funding to assist districts in the PER process, as well as greater involvement of the Minnesota Department of Education through increased technical assistance in the areas of curriculum development and evaluation (Randall, 1989).

One of the most obvious and significant results of the many changes to the original law is the increased state control over curriculum and evaluation decisions made at the local level. Clearly, the intent of the state legislature has been to require "greater accountability and standardization in the area of curriculum and testing" (Randall, 1989, p. 19). At this point in time, however, school districts are only required to report what they are doing in the areas of curriculum development, program planning, and evaluation. They are encouraged, not required, to follow Department of Education guidelines and recommendations in curriculum planning and evaluation procedures. It is speculative as to whether the legislature will continue to increase the amount of control the state has over local programming decisions.

Curriculum Development

Another topic which is important in order to gain a more comprehensive picture of assessment in Minnesota is that of curriculum development. This area also relates to the PER legislation in a significant way. In 1987 the PER law was changed such that the Department of Education was charged with identifying and developing <u>model</u> learner outcomes which represent the basis, or to some the core, of a given curriculum. Such outcomes do not represent the whole curriculum, rather they are viewed as the basic building blocks on which any good curriculum could be based and developed in greater detail. From this set of model outcomes <u>essential</u> learner outcomes are to be identified. "By definiton, model learner outcomes represent skills, knowledge or attitudes which pupils <u>may</u> be exposed to in the curriculum and essential learner outcomes are those skills, knowledge and attitudes that all pupils in Minnesota public schools <u>must</u> have the opportunity to learn" (Randall, 1989, p. 9). The PER law requires that all districts integrate these essential outcomes developed by the Department of Education into their own locally developed curriculum.

The Department of Education is also charged with developing evaluation procedures for these essential learner outcomes. These evaluation procedures are seen as the basis for state level testing, or the Minnesota Assessment of Educational Progress

(MAEP). Thus, it is evident that there is a very close relationship between what is being recommended as the basis for the development of curriculum and what is required to be assessed on a statewide basis. While there is still an emphasis on locally developed curricula some district personnel may feel that the essential learner outcomes represent most of their curricula, thus the distinction between what they feel they <u>can</u> develop versus what they <u>must</u> incorporate from the state level may become blurred.

This interpretation, that the essential outcomes do in fact reflect the bulk of a given curriculum, can yield at least two potentially negative perceptions. First, any educator who strives to teach with the primary goal that their students test well on the state assessment, will be teaching toward the test. This would be unfortunate since the intent is not to develop test driven curricula. A second problem would be the perception that Minnesota has a state mandated curriculum, which it does not, and likely will never have. Since the essential learner outcomes for art education have not been identified yet, due to the fact that the model learner outcomes are just now being developed, avoidance of these potential problems is possible. Efforts will be made to keep those outcomes identified as essential to a minimum.

Aside from the assessment of the essential learner outcomes, curriculum and assessment personnel at the Department of Education are also working on other evaluation procedures. Currently emphasis is being placed on a special project which will offer classroom teachers the opportunity to develop their own tests to assess student performance using state level resources.

The Minnesota Test Item Bank

The Item Bank is a specific assessment project being developed by the Minnesota Department of Education. It is a computer based system which is designed to help educators measure the degree of success that their students achieve in mastering locally developed outcomes as reflected in the curriculum and instruction in their specific school districts. The results of this measurement can be used to evaluate the curriculum and instruction in order to determine where improvements need to be made to increase the degree of student success. For effective evaluation to occur there must be some correspondence between what is taught in the curriculum and the measurement instrument, especially when using a centralized source of test items.

A brief description of the Item Bank in general will be followed by a more detailed description of procedures taken in developing the Visual Art Test Item Bank. The Test Item Bank is a stand alone interactive system on CD-ROM disk, having a capacity of 270,000 items on one disk. It is intended that school districts use IBM-PC technology with a CD ROM drive which will allow direct interactive involvement with the system. A modem can

be used, but it is clumsy and must be operated at 300 or 1200 BAUD. There is an 800 telephone number available to districts to access the system using a modem. At present, the system is in place in 16 test sites around the state funded by the Department of Education, with an additional 7 or 8 districts which have bought the necessary equipment on their own. Using the interactive CD ROM system educators can access the Item Bank to review questions, build a test, and print it on site.

For each subject area an indexing system was developed based on a hierarchical structure of subject, course, outcome, cluster, and subcluster. It was felt that this type of indexing system would facilitate item storage and retrieval better than word search strategies which could get cumbersome with large categories. The indexing system and the program allow teachers to build tests in three basic ways: 1) Request a test by examining items individually; 2) Request a test by computer random sampling; or 3) Review items currently in a test.

All test items in the system are coded to five index levels and one cognitive level. The five index levels are outlined below:

Index Level	Code Description
1) COURSE	For the art item bank the course code is ART
2) SUBJECT	Refers to the instructional level the item is written for: Elementary (K-3, 4-6); Junior High; and Senior High
3) OUTCOME	Refers to the general area of knowledge being assessed: Knowing About Art; Producing Art; and Responding to Art (3 outcomes identified)
4) CLUSTER	Refers to categories which identify more specific areas of knowledge, skills, or affect, e.g. Art in Daily Life; Use of Design Concepts; Aesthetic Qualities (24 identified for the 3 outcomes)
5) SUBCLUSTER	Refers to the most specific level of identification relating to what each test item measures, e.g. Persuasive Functions (SUBCLUSTER) under Art in Daily Life (CLUSTER) The number of SUBCLUSTERS under the CLUSTERS ranges from 4-21, and averages 11.

The last level of classification is referred to as the Cognitive Level and is derived from Bloom's Taxonomy of Educational Objectives (1965). These categories are: Knowing; Applying, Integrating, and Valuing. Following is a description of how these categories relate to the Item Bank:

> If the item calls for the student to know or comprehend; to know specific facts, ideas, or vocabulary and/or to communicate that knowledge (recall) the item would be labeled "K" or Knowing. If the item calls for the student to use learned knowledge or to break down an idea into its parts and/or to perceive the interrelationship between ideas (conceptualization) the item would be labeled "A" or Application. If, on the other hand, the item calls for the student to use knowledge to locate new patterns or relationships or to make decisions or judgements based on given criteria or standards (generalization) the item would be labeled "I" or Integrating. The Valuing classification is not directly related to Bloom's Taxonomy. In this case if the item calls for the student to make a value judgement based on personal criteria rather than stated criteria the item should be labeled "V" or Valuing. (Classification of)

All five index levels and the cognitive level have an alphanumeric code. After an item is keyed into the database the computer assigns it a unique number based on this alphanumeric code. With this code and the assigned number test builders can readily locate any specific item in the Item Bank.

Since the indexing system provides the structure for operating the Item Bank much care went into the development of the art index. The Art Item Bank Index is nearly complete with only one section still being developed. The index is comprised of descriptive words and phrases which represent an overview of the field of art education, categories within which specific test items will be coded and listed. The index reflects the body of knowledge that art educators can draw from to develop learning experiences for their students. It represents the breadth of learning which is possible within the field, not within just one course of study.

A committee of sixteen art educators was formed in the summer of 1988 reflecting balanced choices, geographically (urban and rural districts throughout Minnesota), and in terms of gender and ethnicity. This committee developed the Visual Art Item Bank Index after several months of discussion and deliberation. Though not everyone agreed on all of the terms chosen a consensus was always reached with the realization that the index must be understandable and functional for a variety of educators, some with little or no training in art education, such as elementary generalists.

The next step was to begin securing test items, a process which included purchasing items, using those from the public domain, or otherwise obtaining permission to use items from other sources. The state legislature directed the Department of Education to secure up to ten thousand items per subject area in the Item Bank, a goal which perhaps may be unattainable in any arts area due to the paucity of good tests in the arts. Indeed, obtaining items from other instruments did not prove fruitful. Thus, some process to allow the development of viable art test items for the Item Bank became a necessity. It was decided to contract out the task of writing items to a cadre of educators who would be at least minimally instructed in the fundamentals of test item writing.

Arguments can be raised for and against teacher-prepared versus evaluator-prepared test items. Certainly teachers are much closer to the students and the curriculum, but generally they are ill-prepared in the area of testing and measurement. A typical undergraduate educational psychology course for aspiring teachers covers only the rudiments of test theory, such as central tendency (mean, median, and mode) and variability (standard deviation and variance). Such descriptive statistics are helpful in monitoring student knowledge and performance, however a more in-depth knowledge of testing and measurement is needed for the development and evaluation of effective assessment instruments. Even a knowledge of inferential statistics does not address basic issues in testing and measurement.

Evaluators who have specific training in this area possess the requisite knowledge for item development and test construction. An awareness of item analysis procedures, such as item discrimination and difficulty indices and item characteristic curves, provide evaluators with specific tools for test construction and analysis. Even armed with such tools and skills evaluators may be at a disadvantage due to a lack of familiarity with the curriculum being implemented in a given program. Knowledge of both the curricular offerings and some of the fundamentals of testing and measurement would ensure a closer fit between assessment and instruction. Toward this end, it was felt that art educators familiar with state level curriculum development recommendations could be instructed in techniques of test item construction, in an effort to bridge the gap in abilities between classroom instructor and professional evaluator.

In the fall of 1988 sixteen art educators, many from the original item index committee, attended a conference sponsored by the Department of Education on test item writing. The conference was developed for both art and music educators and was staffed by the author, the music education specialist, and two members of the Assessment Section from the Department of Education. Two intensive days were spent developing skills in the conference participants identifying processes and techniques to write quality test items. Recognizing that two days are not sufficient to develop high level item writing skills it at least represents a

beginning effort to address the problem. These educators were contracted to write 120 items over the next year and submit them to the Department of Education for review and further development if necessary. Efforts have been made to provide follow up assistance for the item writers where needed. A refresher conference will be planned and the process continued into the next year.

Statewide Assessment

Minnesota initiated statewide testing in 1981 with the development of assessment instruments for many disciplines including art education. Three grade levels were assessed in each subject area on a four-year cycle, grades four, eight, and eleven. The assessment program represented an effort to collect statewide data for several purposes:

- to provide valid and reliable data on the knowledge, skills, and attitudes of Minnesota students;
- to provide a means to evaluate curriculum strengths and needs;
- to provide a means to focus developmental activities to improve student learning;
- to measure changes in student performance over time;
- to provide materials, data, and technical assistance to local shool districts engaging in a process of curriculum evaluation through the Local Assessment Option;
- to provide a relevant means for local school districts to meet PER requirements. (Minnesota Statewide)

The original state assessment instruments were based on learner outcome documents and the National Assessment of Educational Progress (NAEP) tests. Several types of comparisons were possible using the data gathered. Minnesota students were compared to national groups and students from individual districts were compared with each other based on the type and size of districts. The state instruments were normed in 1981-82 and renormed in 1985-86.

One of the main purposes of the statewide assessment program was to provide a means to evaluate the strengths and needs of curricula throughout the state, not to report individual student performance. Therefore, a random sample of students were selected from across the state and intact classes were used in order to minimize disruptions to students. In an effort to standardize test administration procedures and reduce reading as a testing variable a pace-tape was used during administration.

This author views the 1981 (Secondary) and 1982 (Elementary) Minnesota assessment instruments as wholly inadequate for many reasons. One very important area of inadequacy with these tests is their inclusion of many attitudinal questions. Items which

elicit an attitudinal response do not have a right or wrong answer. It serves no purpose to compare student scores based on such items from one district with those from another under the assumption that one is comparing what students have learned in an art program. Furthermore, data from items such as these are of little usefulness for educators interested in strengthening art program offerings.

Several changes in the statewide assessment program will be occurring in the next few years. The assessment instruments themselves will be reviewed and completely revised and rewritten during 1990-91. An effort will be made to construct questions which relate more to skills and knowledge that students gain from participation in art programs, rather than how they feel about certain images or objects. The work with the Test Item Bank will aid in the development of items for the state assessment, though it is not anticipated that the makeup of the state instrument will simply be drawn from the Item Bank. Item Bank questions are curriculum level questions having greater specificity than that desired on a state level instrument. Many questions will overlap, however many others on the state test should reflect broader knowledge.

As discussed earlier, once model learner outcomes have been developed for art education essential learner outcomes will be identified. The model outcomes represent the part of a program that students may be exposed to, while the essential outcomes are those that all students must experience. The state test is to be based on these essential learner outcomes. The essential outcomes represent the learning that should be available to all students in Minnesota reagardless of gender, ethnicity, district size, geographic location, or socio-economic status.

After the new state assessment instrument is developed districts will be required to administer it at some point in their curriculum/evaluation cycle. Most likely this will occur during the year that art education is up for review in each district. At any other time if a district wishes to assess in the art area they are free to choose whatever assessment instruments or approaches they desire. Also, even though they must use the state test during their review cycle they do not have to limit themselves to only using the state test. Currently, all districts in the state are required to participate in the state sample for norming purposes once every three years, with the three largest districts (Minneapolis, St. Paul, and Duluth) participating every year.

A recent Department of Education memo summarizes many of the anticipated changes to the statewide assessment program.

1) Testing would be conducted at the most appropriate grade levels for a given discipline rather than establishing fixed grades for all subject areas.

2) Valid statewide sample testing would be conducted at up

to three grade levels per discipline.

3) Testing would be conducted in early spring.

4) Special questionnaires (student, teacher, building) would be developed to accompany each assessment in an attempt to better determine which student, instructional and contextual variables are explanatory of student performance on core curriculum outcomes.

5) Every attempt will be made to utilize alternative kinds of assessment depending upon the measurement requirements for given outcomes.

6) Exploration relative to the feasibility of international comparisons would continue. (Sauter, Erickson, McMillan, & Valdez, 1987)

Summary

Since the PER legislation requires the Department of Education to identify essential learner outcomes which must be incorporated into each district's educational plan some degree of curriculum standardization will result. As long as this standardization is a small part of any district's program it represents only a minimal intrusion into local decision making. Indeed, such standardization tends to be viewed somewhat enthusiastically by art educators who perceive some art programs around the state as being inadequate. On the other hand, some art educators feel they have an excellent program and do not need or desire outside assistance and tend to view any standardization as an unwanted intrusion on their freedom to make all policy decisions locally.

Standardized testing has often been a controversial area and the 1980's is seeing its share of argument and discussion concerning this topic. Though there are questions and arguments as to what standardized testing can offer educators interested in student progress and program development (Hamblen, 1987) the very decision of whether or not to adopt a standardized testing program is often not up to educators. Usually state legislators are the driving force beind such decisions. The decision to administer standardized tests is often made because there is a strong desire to compare students from one district with those in another. The rationale for doing this, however, is not always clear.

Although work with the Test Item Bank is currently based on paper and pencil testing using a multiple choice format a major thrust over the next few years will be to develop non-paper and pencil approaches as well as alternatives to the multiple choice format. The revised state assessment instrument will also reflect this new direction. There is great interest in investigating interactive video as a way to assess actual student performance

rather than the rote knowledge a student possesses. Other areas under investigation include the use of narrative summaries, observational data, self reports and/or journals, reflective interviews, and portfolios of completed and in-process work.

Nationwide, more emphasis needs to be placed on the term evaluation. Many individuals, and especially educators, use the term testing when they very likely mean evaluating. Testing is a subset or one kind of evaluative method at the disposal of educators. It is based on the concept of measurement which involves assigning a number to some characteristic which would otherwise not be considered numerically, such as a behavior, performance, or attitude. Non-testing methods of providing evaluative feedback are absolutely necessary to gain a full appreciation of a student's development and progress in any type of curricular offering. Decisions about evaluation procedures should not be based on the fact that testing procedures are almost always more economical than non-testing approaches, in both human and monetary terms. Such decisions should be based on the desire to obtain the most comprehensive picture possible about student progress.

Since the term accountability often seems threatening to many educators and administrators they often opt for testing approaches because these are seen as more rigorous than non-testing approaches. Efforts are being made to establish more rigorous procedures using observational and narrative data, self reports, and other non-testing approaches (Crane, 1988; Ericsson & Simon, 1984; Howard, 1980; Scott & Hatfield, 1985). A battery of several different evaluation procedures should be used in all school programs. More funding needs to be spent on investigating non-testing approaches. Most importantly, attitudes need to change about the acceptability of all kinds of evaluation. This is true across the educational spectrum including teachers, students, parents, administrators, and school board members.

BIBLIOGRAPHY

Bloom, B. (Ed.). (1965). Taxonomy of educational objectives. New York: David McKay.

Classification of test items. (n.d.). St. Paul, MN: Assessment Section, Minnesota Department of Education.

Crane, J. A. (1988). Evaluation as scientific research. Educational Review, 12, 467-482.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: The Massachusetts Institute of Technology Press.

Hamblen, K. (1987). What general education can tell us about evaluation in art. Studies in Art Education, 28(4), 246-250.

Howard, G. S. (1980). Response-shift bias, a problem in evaluating interventions with pre/post self-reports. Evaluation Review, 4(1), 93-106.

Minnesota educational assessment program: Elementary visual arts. (1982). St. Paul, MN: Assessment Section, Minnesota Department of Education.

Minnesota educational assessment program: Secondary visual arts. (1981). St. Paul, MN: Assessment Section, Minnesota Department of Education.

Minnesota statewide educational assessment program. (n.d.). St. Paul, MN: Assessment Section, Minnesota Department of Education.

National Endowment for the Arts. (1988). Towards civilization: A report on arts education. Washington, D.C.: National Endowment for the Arts.

Randall, R. (1989). Report to the legislature on the Minnesota planning, evaluation and reporting (PER) program. Internal memorandum. St. Paul, MN: Minnesota Department of Education.

Sauter, J., Erickson, W., McMillan, W., & Valdez, G. (1987). Policies for development and implementation of core curriculum, assessment and related processes. Internal memorandum. St. Paul, MN: Minnesota Department of Education.

Scott, M. M., & Hatfield, J. G. (1985). Problems of analyst and observer agreement in naturalistic narrative data. Journal of Educational Measurement, 22(3), 207-218.